

A Gentle Introduction to a Beautiful Theorem of Molien

Holger Schellwat

holger.schellwat@oru.se, Örebro universitet, Sweden
Universidade Eduardo Mondlane, Moçambique

12 January, 2017

Abstract

The purpose of this note is to give an accessible proof of Moliens Theorem in Invariant Theory, in the language of today's Linear Algebra and Group Theory, in order to prevent this beautiful theorem from being forgotten.

Contents

1	Preliminaries	3
2	The Magic Square	6
3	Averaging over the Group	9
4	Eigenvectors and eigenvalues	11
5	Moliens Theorem	13
6	Symbol table	17
7	Lost and found	17
	References	17
	Index	18

Introduction

We present some memories of a visit to the ring zoo in 2004. This time we met an animal looking like a unicorn, known by the name of invariant theory. It is rare, old, and very beautiful. The purpose of this note is to give an almost self contained introduction to and clarify the proof of the amazing theorem of Molien, as presented in [Slo77]. An introduction into this area, and much more, is contained in [Stu93]. There are many very short proofs of this theorem, for instance in [Sta79], [Hu90], and [Tam91].

Informally, Moliens Theorem is a power series generating function formula for counting the dimensions of subrings of homogeneous polynomials of certain degree which are invariant under the action of a finite group acting on the variables. As an appetizer, we display this stunning formula:

$$\Phi_G(\lambda) := \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{id} - \lambda T_g)}$$

We can immediately see elements of linear algebra, representation theory, and enumerative combinatorics in it, all linked together. The paper [Slo77] nicely shows how this method can be applied in Coding theory. For Coding Theory in general, see [Bie04].

Before we can formulate the Theorem, we need to set the stage by looking at some Linear Algebra (see [Rom 08]), Group Theory (see [Hu96]), and Representation Theory (see [Sag 91] and [Tam91]).

1 Preliminaries

Let $V \cong \mathbf{C}^n$ be a finite dimensional complex inner product space with orthonormal basis $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ and let $\mathbf{x} = (x_1, \dots, x_n)$ be the orthonormal basis of the algebraic dual space V^* satisfying $\forall 1 \leq i, j \leq n : x_i(\mathbf{e}_j) = \delta_{ij}$. Let G be a finite group acting unitarily linear on V from the left, that is, for every $g \in G$ the mapping $V \rightarrow V, \mathbf{v} \mapsto g \cdot \mathbf{v}$ is a unitary bijective linear transformation. Using coordinates, this can be expressed as $[g \cdot \mathbf{v}]_{\mathcal{B}} = [g]_{\mathcal{B}, \mathcal{B}} [\mathbf{v}]_{\mathcal{B}}$, where $[g]_{\mathcal{B}, \mathcal{B}}$ is unitary. Thus, the action is a unitary representation of G , or in other words, a G -module. Note that we are using left composition and column vectors, i.e. $\mathbf{v} = (v_1, \dots, v_n) \stackrel{\text{convention}}{=} [v_1 \ v_2 \ \dots \ v_n]^\top$, c. f. [Ant73].

The elements of V^* are linear forms (linear functionals), and the elements x_1, \dots, x_n , looking like variables, are also linear forms, this will be important later.

Thinking of x_1, \dots, x_n as variables, we may view (see [Tam91]) $S(V^*)$, the *symmetric algebra* on V^* as the algebra $R := \mathbf{C}[\mathbf{x}] := \mathbf{C}[x_1, \dots, x_n]$ of polynomial functions $V \rightarrow \mathbf{C}$ or polynomials in these variables (linear forms). It is naturally graded by degree as $R = \bigoplus_{d \in \mathbf{N}} R_d$, where R_d is the vector space spanned by the polynomials of (total) degree d , in particular, $R_0 = \mathbf{C}$, and $R_1 = V^*$.

The action of G on V can be lifted to an action on R .

1.1 Proposition. *Let V, G, R as above. Then the mapping $\cdot : G \times R \rightarrow R, (g, f) \mapsto g \cdot f$ defined by $(g \cdot f)(\mathbf{v}) := f(g^{-1} \cdot \mathbf{v})$ for $\mathbf{v} \in V$ is a left action.*

Proof. For $\mathbf{v} \in V, g, h \in G$, and $f \in R$ we check

1. $(1 \cdot f)(\mathbf{v}) = f(1^{-1} \cdot \mathbf{v}) = f(1 \cdot \mathbf{v}) = f(\mathbf{v})$
- 2.

$$\begin{aligned} ((hg) \cdot f)(\mathbf{v}) &= f((hg)^{-1} \cdot \mathbf{v}) = f((g^{-1}h^{-1}) \cdot \mathbf{v}) = \\ &= f(g^{-1} \cdot (h^{-1} \cdot \mathbf{v})) = (g \cdot f)(h^{-1} \cdot \mathbf{v}) = (h \cdot (g \cdot f))(\mathbf{v}) \end{aligned}$$

□

In fact, we know more.

1.2 Proposition. *Let V, G, R as above. For every $g \in G$, the mapping $T_g : R \rightarrow R, f \mapsto g \cdot f$ is an algebra automorphism preserving the grading, i.e. $g \cdot R_d \subset R_d$ (here we do not bother about surjectivity).*

Proof. For $\mathbf{v} \in V, g \in G, c \in \mathbf{C}$, and $f, f' \in R$ we check

- 1.

$$\begin{aligned} (g \cdot (f + f'))(\mathbf{v}) &= (f + f')(g^{-1} \cdot \mathbf{v}) = f(g^{-1} \cdot \mathbf{v}) + f'(g^{-1} \cdot \mathbf{v}) = \\ &= (g \cdot f)(\mathbf{v}) + (g \cdot f')(\mathbf{v}) = (g \cdot (f + f'))(\mathbf{v}), \text{ thus } g \cdot (f + f') = g \cdot f + g \cdot f' \end{aligned}$$

- 2.

$$\begin{aligned} (g \cdot (f \cdot f'))(\mathbf{v}) &= (f \cdot f')(g^{-1} \cdot \mathbf{v}) = f(g^{-1} \cdot \mathbf{v}) \cdot f'(g^{-1} \cdot \mathbf{v}) = \\ &= (g \cdot f)(\mathbf{v}) \cdot (g \cdot f')(\mathbf{v}) = (g \cdot (f \cdot f'))(\mathbf{v}), \text{ thus } g \cdot (f \cdot f') = g \cdot f \cdot g \cdot f' \end{aligned}$$

3. $(g.(cf))(\mathbf{v}) = (cf)(g^{-1}.\mathbf{v}) = c(f(g^{-1}.\mathbf{v})) = c((g.f)(\mathbf{v})) = (c(g.f))(\mathbf{v})$
4. By part 2. it is clear that the grading is preserved.
5. To show that $f \mapsto g.f$ is bijective it is enough to show that this mapping is injective on the finite dimensional homogeneous components R_d . Let us introduce a name for this mapping, say $T_g^d : R_d \rightarrow R_d, f \mapsto g.f$. Now $f \in \ker(T_g^d)$ implies that $g.f = 0 \in R_d$, i.e. $g.f$ is a polynomial mapping from V to \mathbf{C} of degree d vanishing identically, $\forall \mathbf{v} \in V : (g.f)(\mathbf{v}) = 0$. By definition of the extended action we have $\forall \mathbf{v} \in V : f(g^{-1}.\mathbf{v}) = 0$. Since G acts on V this implies that $\forall \mathbf{v} \in V : f(\mathbf{v}) = 0$, so f is the zero mapping. Since our ground field has characteristic 0, this implies that f is the zero polynomial, which we may view as an element of every R_d . See for instance [Cox91], proposition 5 in section 1.1.
6. Note that every T_g^d is also surjective, since all group elements have their inverse in G .

□

Both propositions together give us a homomorphism from G into $\text{Aut}(R)$. They also clarify the rôle of the *induced* matrices, which are classical in this area, as mentioned in [Slo77]. Since the monomials x_1, \dots, x_n of degree one form a basis for R_1 , it follows from the proposition that their products $\mathbf{x}_2 := (x_1^2, x_1x_2, x_1x_3, \dots, x_1x_n, x_2^2, x_2x_3, \dots)$ form a basis for R_2 , and, in general, the monomials of degree d in the linear forms (!) x_1, \dots, x_n form a basis \mathbf{x}_d of R_d . Clearly, they certainly span R_d , and by the last observation in the last proof they are linearly independent.

1.3 Definition. In the context from above, that is $g \in G$, $f \in R^d$, and $\mathbf{v} \in V$, we define

$$T_g^d : R_d \rightarrow R_d, f \mapsto g.f : R^d \rightarrow \mathbf{C}, \mathbf{v} \mapsto f(g^{-1}.\mathbf{v}) = f(T_{g^{-1}}(\mathbf{v})).$$

1.4 Remark. In particular, we have $(T_g^1(f))(\mathbf{v}) = f(T_{g^{-1}}(\mathbf{v}))$, see proposition 1.6 below.

Keep in mind that a function $f \in R_d$ maps to $T_g^d(f) = g.f$. Setting $A_g := [T_g^1]_{\mathbf{x}, \mathbf{x}}$, then $A_g^{[d]} := [T_g^d]_{\mathbf{x}_d, \mathbf{x}_d}$ is the d -th induced matrix in [Slo77], because $T_g^1(f \cdot f') = T_g^1(f) \cdot T_g^1(f')$. Also, if f, f' are eigenvectors of T_g^1 corresponding to the eigenvalues λ, λ' , then $f \cdot f'$ is an eigenvector of T_g^2 with eigenvalue $\lambda \cdot \lambda'$, because $T_g(f \cdot f') = T_g(f) \cdot T_g(f') = (\lambda f) \cdot (\lambda' f') = (\lambda \lambda')(f \cdot f')$. All this generalizes to $d > 2$, we will get back to that later.

We end this section by verifying two little facts needed in the next section.

1.5 Proposition. *The first induced operator of the inverse of a group element $g \in G$ is given by $T_{g^{-1}}^1 = (T_g^1)^{-1}$.*

Proof. Since $\dim(V^*) < \infty$, it is sufficient to prove that $T_{g^{-1}}^1 \circ T_g^1 = \text{id}_{V^*}$. Keep in mind that $(T_g^1(f))(\mathbf{v}) = f(T_{g^{-1}}(\mathbf{v}))$. For arbitrary $f \in V^*$ we see that

$$(T_{g^{-1}}^1 \circ T_g^1)(f) = T_{g^{-1}}^1(T_g^1(f)) = T_{g^{-1}}^1(g.f) = g^{-1}.(g.f) = (g^{-1}g).f = f.$$

□

We will be mixing group action notation and composition freely, depending on the context. The following observation is a translation device.

1.6 Proposition. *For $g \in G$ and $f \in V^*$ the following holds:*

$$T^1(f) = g.f = f \circ T_{g^{-1}}.$$

Proof. For $\mathbf{v} \in V$ we see $(T^1(f))(\mathbf{v}) = (g.f)(\mathbf{v}) \stackrel{\text{def}}{=} f(g^{-1}.\mathbf{v}) = f(T_{g^{-1}}(\mathbf{v}))$. \square

2 The Magic Square

Remember that we require a unitary representation of G , that is the operators $T_g : V \rightarrow V$ need to be unitary, i.e. $\forall g \in G : (T_g)^{-1} = (T_g)^*$. The first goal of this sections is to show that this implies that the induced operators $T_g^d : R_d \rightarrow R_d, f \mapsto g \cdot f$ are also unitary. We saw that $T_g^1 = V^*$, the algebraic dual of V . In order to understand the operator duals of V and V^* we need to look on their inner products first. We may assume that the operators T_g are unitary with respect to the standard inner product $\langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_{\mathcal{B}, \mathcal{B}} \bullet \overline{[\mathbf{v}]_{\mathcal{B}, \mathcal{B}}}$, where \bullet denotes the dot product.

Before we can speak of unitarity of the induced operators T_g^d we have to make clear which inner product applies on $R^1 = V^*$. Quite naively, for $f, g \in V^*$ we are tempted to define $\langle f, g \rangle = [f]_{\mathbf{x}, \mathbf{x}} \bullet \overline{[g]_{\mathbf{x}, \mathbf{x}}}$.

We will motivate this in a while, but first we take a look at the diagram in [Rom 08], chapter10, with our objects:

$$\begin{array}{ccc}
 & \xleftarrow{T_g^\times} & \\
 R^1 = V^* & \xrightarrow{T_g^1} & V^* = R^1 \\
 \downarrow P & & \downarrow P \\
 V & \xrightarrow{T_g} & V \\
 & \xleftarrow{T_g^*} &
 \end{array}$$

Here P (“Rho”) denotes the Riesz map, see [Rom 08], Theorem 9.18, where it is called R , but R denotes already our big ring. We started by looking at the operator T_g , which is unitary, so its inverse is the Hilbert space adjoint T_g^* . Omitting the names of the bases we have $[T_g^*] = [T_g]^*$. We also see the operator adjoint T_g^\times with matrix $[T_g^\times] = [T_g]^\top$, the transpose. However, the arrow for T_g^1 is not in the original diagram, but soon we will see it there, too.

Fortunately, the Riesz map P turns a linear form into a vector and its inverse $\tau : V \rightarrow V^*$ maps a vector to a linear form, both are conjugate isomorphisms. This is mostly all we need in order to show that T_g^1 is unitary. In the following three propositions we use that V has the orthonormal basis \mathcal{B} and that V^* has the orthonormal basis \mathbf{x} .

2.1 Proposition. *For every $f \in V^*$ the coordinates of its Riesz vector are given by*

$$[P(f)]_{\mathbf{e}} = (\overline{f(\mathbf{e}_1)}, \dots, \overline{f(\mathbf{e}_n)}).$$

Proof. Writing τ for the inverse of P , we need to show that

$$P(f) = \sum_{i=1}^n \overline{f(\mathbf{e}_i)} \mathbf{e}_i$$

which is equivalent to

$$f = \tau \left(\sum_{i=1}^n \overline{f(\mathbf{e}_i)} \mathbf{e}_i \right).$$

It is sufficient to show the latter for values of f on the basis vectors \mathbf{e}_j , $1 \leq j \leq n$. We obtain

$$\begin{aligned} \left(\tau \left(\sum_{i=1}^n \overline{f(\mathbf{e}_i)} \mathbf{e}_i \right) \right) (\mathbf{e}_j) &= \left\langle \mathbf{e}_j, \left(\sum_{i=1}^n \overline{f(\mathbf{e}_i)} \mathbf{e}_i \right) \right\rangle = \sum_{i=1}^n \left\langle \mathbf{e}_j, \left(\overline{f(\mathbf{e}_i)} \mathbf{e}_i \right) \right\rangle \\ &= \overline{f(\mathbf{e}_i)} \sum_{i=1}^n \langle \mathbf{e}_j, \mathbf{e}_i \rangle = f(\mathbf{e}_i) \cdot 1. \end{aligned}$$

□

In particular, this implies that $P(x_i) = \mathbf{e}_i$.

2.2 Proposition. *Our makeshift inner product on V^* satisfies*

$$\langle f, g \rangle = \langle P(f), P(g) \rangle,$$

where $f, g \in V^*$.

Proof. By our vague definition we have $\langle f, g \rangle = [f]_{\mathbf{x}, \mathbf{x}} \bullet \overline{[g]_{\mathbf{x}, \mathbf{x}}}$. It is enough to show that $\langle x_i, x_j \rangle = \langle P(x_i), P(x_j) \rangle$. From the comment after the proof of Proposition 2.1 we obtain

$$\langle P(x_i), P(x_j) \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j = [x_i]_{\mathbf{x}, \mathbf{x}} \bullet \overline{[x_j]_{\mathbf{x}, \mathbf{x}}}.$$

□

Hence, our guess for the inner product on V^* was correct. We will now relate the Riesz vector of $f \in V^*$ to the Riesz vector of $f \circ T_g^{-1}$. Recall that the Riesz vector of $f \in V^*$ is the unique vector $\mathbf{w} = P(f)$ such that $f(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{v} \in V$. If $f \neq 0$ it can be found by scaling any nonzero vector in the cokernel of f , which is one-dimensional, see [Rom 08], in particular Theorem 9.18.

2.3 Proposition. *Let $T_g : V \rightarrow V$ be unitary, $f \in V^*$, $\mathbf{w} = P(f)$ the vector of $f \in V^*$. Then $T_g(\mathbf{w})$ is the Riesz vector of $f \circ T_g^{-1}$, i.e. the Riesz vector of $T_g^1(f)$.*

Proof. We may assume that $f \neq 0$. Using the notation $\langle \mathbf{w} \rangle$ for the one-dimensional subspace spanned by \mathbf{w} , we start with a little diagram:

$$\langle \mathbf{w} \rangle \odot \ker(f) \xrightarrow{T_g} \langle T_g(\mathbf{w}) \rangle \odot \ker(f \circ T_g^{-1}),$$

where \odot denotes the orthogonal direct sum.

We need to show that $f \circ T_g^{-1} = \langle \cdot, T_g(\mathbf{w}) \rangle$, i.e. that $(f \circ T_g^{-1})(\mathbf{v}) = \langle \mathbf{v}, T_g(\mathbf{w}) \rangle$ for all $\mathbf{v} \in V$. Since $\mathbf{w} = P(f)$ the vector of f , we have $f(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{v} \in V$. We obtain

$$(f \circ T_g^{-1})(\mathbf{v}) = \langle T_g^{-1}(\mathbf{v}), \mathbf{w} \rangle \stackrel{T_g \text{ unitary}}{=} \langle \mathbf{v}, T_g(\mathbf{w}) \rangle.$$

From remark 1.4 we conclude that $f \circ T_g^{-1} = T_g^1(f)$.

□

Observe that proposition 2.3 implies the commutativity of the following two diagrams.

$$\begin{array}{ccc} V^* & \xrightarrow{T_g^1} & V^* \\ \downarrow P & & \downarrow P \\ V & \xrightarrow{T_g} & V \end{array} \quad \text{and} \quad \begin{array}{ccc} V^* & \xrightarrow{(T_g^1)^{-1}} & V^* \\ \downarrow P & & \downarrow P \\ V & \xrightarrow{(T_g)^{-1}} & V \end{array}$$

Indeed, 2.3 implies

$$P \circ T_g^1 = T_g \circ P \quad (1)$$

$$P \circ (T_g^1)^{-1} = (T_g)^{-1} \circ P \quad (2)$$

2.4 Proposition. *The first induced operator T_g^1 is unitary.*

Proof. We may use that T_g is unitary, that is,

$$\langle T_g(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, (T_g)^{-1}(\mathbf{w}) \rangle = \langle \mathbf{v}, (T_{g^{-1}})(\mathbf{w}) \rangle \quad (*).$$

Let $f, h \in V^*$ arbitrary, $\mathbf{w} := P(f)$, and $\mathbf{u} := P(h)$. We need to check that $\langle (T_g^1)(f), h \rangle = \langle f, (T_g^1)^{-1}(h) \rangle$. We see that

$$\begin{aligned} \langle (T_g^1)(f), h \rangle &\stackrel{\text{proposition 2.2}}{=} \langle (P \circ T_g^1)(f), P(h) \rangle \stackrel{(1)}{=} \langle (T_g \circ P)(f), P(h) \rangle \\ &= \langle (T_g(P))(f), P(h) \rangle = \langle T_g(\mathbf{w}), \mathbf{u} \rangle \stackrel{*}{=} \langle \mathbf{w}, T_g^{-1}(\mathbf{u}) \rangle \\ &= \langle P(f), T_g^{-1}(P(h)) \rangle = \langle P(f), (T_g^{-1} \circ P)(h) \rangle \\ &\stackrel{(2)}{=} \langle P(f), (P \circ (T_g^1)^{-1})(h) \rangle = \langle P(f), P((T_g^1)^{-1}(h)) \rangle \\ &= \langle f, (T_g^1)^{-1}(h) \rangle \end{aligned}$$

□

After having looked at eigenvalues we will see that this generalizes to higher degree, that T_g^d is diagonalizable for all $d \in \mathbf{Z}^+$. But first let us look at the matrix version of proposition 2.4.

2.5 Proposition.

$$[T_g^1]_{\mathbf{x}, \mathbf{x}} = \overline{[T_g]_{\mathbf{e}, \mathbf{e}}}$$

Proof. Let $A := [T_g]_{\mathcal{B}, \mathcal{B}} = [A_1 | \cdots | A_i | \cdots | A_n] = [a_{i,j}]$ and $B := [T_g^1]_{\mathbf{x}, \mathbf{x}} = [B_1 | \cdots | B_i | \cdots | B_n] = [b_{i,j}]$. We will use the commutativity of the diagram, i.e. $P^{-1} \circ T_g \circ P = T_g$, which we will mark as □. No, the proof is not finished here. We get $T_g(\mathbf{e}_i) = A_i = \sum_{k=1}^n a_{k,i} \mathbf{e}_k$ and

$$\begin{aligned} T_g^1(x_i) &\stackrel{\square}{=} (P^{-1} \circ T_g \circ P)(x_i) = P^{-1}(T_g(P(x_i))) \\ &\stackrel{2.1}{=} P^{-1}(T_g(\mathbf{e}_i)) = P^{-1} \left(\sum_{k=1}^n a_{k,i} \mathbf{e}_k \right) \stackrel{\text{konj.}}{=} \sum_{k=1}^n \overline{a_{k,i}} P^{-1}(\mathbf{e}_k) \\ &\stackrel{2.1}{=} \sum_{k=1}^n \overline{a_{k,i}} x_k \end{aligned}$$

On the other hand, $[T_g^1(x_i)]_{\mathbf{x}} = [T_g^1]_{\mathbf{x}, \mathbf{x}} \mathbf{e}_i = B_i$ implies $T_g^1(x_i) = \sum_{k=1}^n b_{k,i} \mathbf{e}_k$. Together we obtain $b_{k,i} = \overline{a_{k,i}}$, and the proposition follows. □

3 Averaging over the Group

Now we apply averaging to obtain self-adjoint operators.

3.1 Definition. We define the following operators:

1. $\hat{T} : V \rightarrow V, \mathbf{v} \mapsto \hat{T}(\mathbf{v}) := \frac{1}{|G|} \sum_{g \in G} T_g(\mathbf{v})$
2. $\hat{T}^1 : V^* \rightarrow V^*, f \mapsto \hat{T}^1(f) := \frac{1}{|G|} \sum_{g \in G} T_g^1(f)$

These are sometimes called the *Reynolds* operator of G .

3.2 Proposition. *The operators \hat{T} and \hat{T}^1 are self-adjoint (Hermitian).*

Proof. The idea of the averaging trick is that if $g \in G$ runs through all group element and $g' \in G$ is fixed, then the products $g'g$ run also through all group elements. We will make use of the facts that every T_g and every T_g^1 is unitary.

1. We need to show that $\langle \hat{T}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \hat{T}(\mathbf{w}) \rangle$ for arbitrary $\mathbf{v}, \mathbf{w} \in V$.
We obtain

$$\begin{aligned} \langle \hat{T}(\mathbf{v}), \mathbf{w} \rangle &= \left\langle \frac{1}{|G|} \sum_{g \in G} T_g(\mathbf{v}), \mathbf{w} \right\rangle = \frac{1}{|G|} \sum_{g \in G} \langle T_g(\mathbf{v}), \mathbf{w} \rangle \\ &\stackrel{\text{unit.}}{=} \frac{1}{|G|} \sum_{g \in G} \langle \mathbf{v}, (T_g)^{-1}(\mathbf{w}) \rangle = \frac{1}{|G|} \sum_{g \in G} \langle \mathbf{v}, (T_{g^{-1}})(\mathbf{w}) \rangle \\ &= \frac{1}{|G|} \sum_{g' \in G} \langle \mathbf{v}, (T_{g'})(\mathbf{w}) \rangle = \langle \mathbf{v}, \hat{T}(\mathbf{w}) \rangle \end{aligned}$$

2. The same proof, *mutatis mutandis*, replacing $\hat{T} \leftrightarrow \hat{T}^1$, $T_g \leftrightarrow T_g^1$, $\mathbf{v} \leftrightarrow f$, and $\mathbf{w} \leftrightarrow h$ shows that $\langle \hat{T}^1(f), h \rangle = \langle f, \hat{T}^1(h) \rangle$.

□

Consequently, \hat{T} and \hat{T}^1 are unitarily diagonalizable with real spectrum.

3.3 Proposition. *The operators \hat{T} and \hat{T}^1 are idempotent, i.e.*

1. $\hat{T} \circ \hat{T} = \hat{T}$
2. $\hat{T}^1 \circ \hat{T}^1 = \hat{T}^1$.

In particular, the eigenvalues of both operators are either 0 or 1.

Proof. Again, we show only one part, the other part is analog. To begin with, let $s \in G$ be fixed. Then

$$\begin{aligned} T_s \circ \hat{T} &= T_s \circ \frac{1}{|G|} \sum_{g \in G} T_g = \frac{1}{|G|} \sum_{g \in G} T_s \circ T_g \\ &= \frac{1}{|G|} \sum_{g \in G} T_{sg} = \frac{1}{|G|} \sum_{g' \in G} T_{g'} = \hat{T}. \end{aligned}$$

From this it follows that

$$\begin{aligned}\hat{T} \circ \hat{T} &= \left(\frac{1}{|G|} \sum_{g \in G} T_g \right) \circ \hat{T} = \frac{1}{|G|} \sum_{g \in G} T_g \circ \hat{T} \stackrel{\text{above}}{=} \frac{1}{|G|} \sum_{g \in G} \hat{T} \\ &= \frac{1}{|G|} \cdot |G| \cdot \hat{T} = \hat{T}.\end{aligned}$$

From $\hat{T} \circ \hat{T} = \hat{T}$ we conclude that $\hat{T} \circ (\hat{T} - \text{id}) = 0$. Thus the minimal polynomial of T divides the polynomial $\lambda(\lambda - 1)$, so all eigenvalues are contained in $\{0, 1\}$. \square

We will now look at the eigenvalues of T_g and T_g^1 and their interrelation. Since both operators are unitary, their eigenvalues have absolute value 1.

- 3.4 Proposition.** 1. If $\mathbf{v} \in V$ is an eigenvector of T_g for the eigenvalue λ , then \mathbf{v} is an eigenvector of $T_{g^{-1}}$ for the eigenvalue $\bar{\lambda} = \frac{1}{\lambda}$.
2. If $f \in V^*$ is an eigenvector of T_g^1 for the eigenvalue λ , then f is an eigenvector of $T_{g^{-1}}^1$ for the eigenvalue $\frac{1}{\lambda}$.
3. If $f \in V^*$ is an eigenvector of T_g^1 for the eigenvalue λ , then $P(f) \in V$ is an eigenvector of T_g for the eigenvalue $\bar{\lambda} = \frac{1}{\lambda}$.
4. If $\mathbf{v} \in V$ is an eigenvector of T_g for the eigenvalue λ , then $P^{-1}(\mathbf{v}) \in V^*$ is an eigenvector of T_g^1 for the eigenvalue $\bar{\lambda} = \frac{1}{\lambda}$.

Proof. We will make use of the commutativity of Proposition 2.3. Observe that $g \cdot \mathbf{v} = T_g(\mathbf{v})$ and $g \cdot f = f \circ T_g$.

1.

$$\begin{aligned}T_g(\mathbf{v}) = g \cdot \mathbf{v} = \lambda \mathbf{v} &\implies g^{-1} \cdot g \cdot \mathbf{v} = g^{-1} \cdot \lambda \mathbf{v} \implies g^{-1} \cdot g \cdot \mathbf{v} = \lambda g^{-1} \cdot \mathbf{v} \\ &\implies \mathbf{v} = \lambda g^{-1} \cdot \mathbf{v} \implies T_{g^{-1}}(\mathbf{v}) = g^{-1} \cdot \mathbf{v} = \frac{1}{\lambda} \mathbf{v}\end{aligned}$$

2.

$$\begin{aligned}T_g^1(f) = g \cdot f = \lambda f &\implies g^{-1} \cdot g \cdot f = g^{-1} \cdot \lambda f \implies g^{-1} \cdot g \cdot f = \lambda g^{-1} \cdot f \\ &\implies f = \lambda g^{-1} \cdot f \implies T_{g^{-1}}^1(f) = g^{-1} \cdot f = \frac{1}{\lambda} f\end{aligned}$$

3.

$$\begin{aligned}T_g^1(f) = \lambda f &\stackrel{P \circ}{\implies} P(T_g^1(f)) = P(\lambda f) \stackrel{(1)}{\implies} T_g(P(f)) = P(\lambda f) \\ &\implies T_g(P(f)) = \bar{\lambda} P(f) = \frac{1}{\lambda} P(f)\end{aligned}$$

4.

$$\begin{aligned}T_g(\mathbf{v}) = \lambda \mathbf{v} &\stackrel{P^{-1} \circ}{\implies} P^{-1}(T_g(\mathbf{v})) = P^{-1}(\lambda \mathbf{v}) \stackrel{\square}{\implies} (T_g^1 \circ P^{-1})(\mathbf{v}) = \bar{\lambda} P^{-1}(\mathbf{v}) \\ &\implies T_g^1(P^{-1}(\mathbf{v})) = \frac{1}{\lambda} P^{-1}(\mathbf{v})\end{aligned}$$

\square

This implies that if we consider the union of the spectra over all $g \in G$, then we obtain the same (multi)set, no matter if we take T_g or T_g^1 .

4 Eigenvectors and eigenvalues

Now we continue from where we left at the end of section 1, fixing one group element $g \in G$ and compare T_g^1 with T_g^d for $d > 1$. By a method called *stars and bars* it is easy to see that

$$\tilde{d} := \dim_{\mathbf{C}}(R_d) = \frac{(n+d+1)!}{(n-1)!d!}.$$

Remember that every T_g^1 is unitarily diagonalizable with eigenvalues of absolute value 1. If $\text{spec}(T_g^1) = (\omega_1, \dots, \omega_n) \in U(1)^n$, then V^* has an orthonormal basis $\mathbf{y}_g^1 := (y_1, \dots, y_n)$, such that $T_g^1(y_i) = \omega_i \cdot y_i$ for all $1 \leq i \leq n$, and $[T_g^1]_{\mathbf{y}_g^1, \mathbf{y}_g^1} = \text{diag}(\omega_1, \dots, \omega_n)$. Moreover,

$$[T_g^1]_{\mathbf{y}_g^1, \mathbf{y}_g^1} = [\text{id}]_{\mathbf{y}_g^1, \mathbf{x}} \cdot [T_g^1]_{\mathbf{x}, \mathbf{x}} \cdot [\text{id}]_{\mathbf{x}, \mathbf{y}_g^1} = \text{diag}(\omega_1, \dots, \omega_n),$$

where $[\text{id}]_{\mathbf{y}_g^1, \mathbf{x}} = [\text{id}]_{\mathbf{x}, \mathbf{y}_g^1}^*$ is unitary.

For $d > 1$ put

$$\mathbf{x}^d := (x_1^d, x_2^d, \dots, x_n^d, x_1^{d-1}x_2, x_1^{d-1}x_3, \dots, x_1^{d-1}x_n, \dots) =: (\tilde{x}_1, \dots, \tilde{x}_{\tilde{d}}),$$

all monomials in the x_i of total degree d , numbered from 1 to \tilde{d} .

These are certainly linear independent, since we have no relations amongst the variables, and span R_d , since every monomial of total degree d can be written as a linear combination of these. So they form a basis for R_d . We will not require that this can be made into an orthonormal basis, we do not even consider any inner product on R_d for $d > 1$.

We rather want to establish that

$$\mathbf{y}^d := (y_1^d, y_2^d, \dots, y_n^d, y_1^{d-1}y_2, y_1^{d-1}y_3, \dots, y_1^{d-1}y_n, \dots) =: (\tilde{y}_1, \dots, \tilde{y}_{\tilde{d}})$$

is a basis of eigenvectors of T_g^d diagonalizing T_g^d , using the same numbering.

Arranging the eigenvalues of T_g^1 in the same way we put

$$\omega^d := (\omega_1^d, \omega_2^d, \dots, \omega_n^d, \omega_1^{d-1}\omega_2, \omega_1^{d-1}\omega_3, \dots, \omega_1^{d-1}\omega_n, \dots) =: (\tilde{\omega}_1, \dots, \tilde{\omega}_{\tilde{d}}).$$

Now we establish that the \tilde{y}_i , $1 \leq i \leq \tilde{d}$ are the eigenvectors for the eigenvalues $\tilde{\omega}_i$ of T_g^d .

4.1 Proposition. *In the context above,*

$$T_g^d(\tilde{y}_i) = \tilde{\omega}_i \cdot \tilde{y}_i$$

for all $1 \leq i \leq \tilde{d}$.

Proof. The key is proposition 1.2, as in the preliminary observations at the end of section 1. Let

$$\tilde{y}_i = \prod_{j=1}^n y_j^{\epsilon_j^i}$$

and

$$\tilde{\omega}_i = \prod_{j=1}^n \omega_j^{\epsilon_j^i},$$

where $\epsilon_j \in \mathbf{N}$ and the sum of these exponents is d . Then

$$T_g^d(\tilde{y}_i) = T_g^d \left(\prod_{j=1}^n y_j^{\epsilon_j} \right) = \prod_{j=1}^n T_g^1(y_j^{\epsilon_j}) = \prod_{j=1}^n \omega_j^{\epsilon_j} y_j^{\epsilon_j} = \tilde{\omega}_i \cdot \tilde{y}_i$$

□

As a consequence, R_d has a basis of eigenvectors of T_g^d and T_g^d is similar to the diagonal matrix $\text{diag}(\tilde{\omega}_1, \dots, \tilde{\omega}_{\tilde{d}})$.

5 Moliens Theorem

We will now make some final preparations and then present the proof of Moliens Theorem.

For $f \in R$ and $g \in G$ we say that f is an *invariant* of g if $g.f = f$ and that f is a (simple) invariant of G if $\forall g \in G : g.f = f$. The method of averaging from section 3 can also be applied to create invariants:

5.1 Proposition. *For $f \in V^*$ put $\hat{f} := \hat{T}^1(f)$. Then \hat{f} is an invariant of G .*

Proof. Let $g \in G$ be arbitrary. We will show that $g.\hat{f} = \hat{f}$. Clearly, from proposition 1.6 we get that

$$\begin{aligned} g.\hat{f} &= \hat{f} \circ T_{g^{-1}} = (\hat{T}^1(f)) \circ T_{g^{-1}} \\ &= \left(\frac{1}{|G|} \sum_{s \in G} T_s^1(f) \right) \circ T_{g^{-1}} = \left(\frac{1}{|G|} \sum_{s \in G} f \circ T_{s^{-1}} \right) \circ T_{g^{-1}} \\ &= \frac{1}{|G|} \sum_{s \in G} f \circ T_{s^{-1}} \circ T_{g^{-1}} = \frac{1}{|G|} \sum_{t \in G} f \circ T_{t^{-1}} = \hat{f}. \end{aligned}$$

□

Now, we call

$$R^G := \{ f \in R : \forall g \in G : g.f = f \}$$

the *algebra of invariants* of G .

5.2 Proposition. *R^G is a subalgebra of R .*

Proof. Since the mapping $f \mapsto g.f$ is linear for every $g \in G$, R^G is the intersection of subspaces, and hence a subspace. Let us check the subring conditions in more detail. For arbitrary $g \in G$, $f, h \in R^G$, and $\mathbf{v} \in V$ we have $g.f = f$, $g.h = h$

1. For the zero $0 \in R$ we obtain $(g.0)(\mathbf{v}) = 0(g^{-1}.\mathbf{v}) = 0(\mathbf{v})$, so $0 \in R^G$.
2. We see

$$\begin{aligned} g.(f - h)(\mathbf{v}) &= (f - h)(g^{-1}.\mathbf{v}) = f(g^{-1}.\mathbf{v}) - h(g^{-1}.\mathbf{v}) \\ &= (g.f)(\mathbf{v}) - (g.h)(\mathbf{v}) = f(\mathbf{v}) - h(\mathbf{v}) = (f - h)(\mathbf{v}) \end{aligned}$$

3. Likewise,

$$\begin{aligned} g.(f \cdot h)(\mathbf{v}) &= (f \cdot h)(g^{-1}.\mathbf{v}) = f(g^{-1}.\mathbf{v}) \cdot h(g^{-1}.\mathbf{v}) \\ &= (g.f)(\mathbf{v}) \cdot (g.h)(\mathbf{v}) = f(\mathbf{v}) \cdot h(\mathbf{v}) = (f \cdot h)(\mathbf{v}). \end{aligned}$$

□

Our subalgebra R^G is graded in the same way as R .

5.3 Proposition. *The algebra of invariants of G is naturally graded as*

$$R^G = \bigoplus_{d \in \mathbf{N}} R_d^G,$$

where $R_d^G = \{ f \in R_d : \forall g \in G : g.f = f \}$, called the d -th homogeneous component of R^G .

Proof. This follows directly from proposition 1.1 and proposition 1.2. \square

5.4 Definition (Molien series). Viewing R_d^G as a vector space, we define

$$a_d := \dim_{\mathbb{C}} R_d^G,$$

the number of linearly independent homogeneous invariants of degree $d \in \mathbb{N}$, and

$$\Phi_G(\lambda) := \sum_{d \in \mathbb{N}} a_d \lambda^d,$$

the *Molien series* of G .

Thus, the Molien series of G is an ordinary power series generating function whose coefficients are the numbers of linearly independent homogeneous invariants of degree d . The following beautiful formula gives these numbers, its proof is the aim of this paper.

5.5 Theorem (Molien, 1897).

$$\Phi_G(\lambda) := \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{id} - \lambda T_g)}$$

Following [Slo77] we first look the number a_1 of linearly independent homogeneous invariants of degree d .

5.6 Theorem (Theorem 13 in [Slo77]).

$$a_1 = \text{Tr}(\hat{T}) = \text{Tr}(\hat{T}^1)$$

Proof. First, we note that the equation $\text{Tr}(\hat{T}) = \text{Tr}(\hat{T}^1)$ follows from the remark at the end of section 3, since the sum for the trace runs over all group elements. Remember that the trace is independent of the choice of basis. From proposition 3.3 we know that both operators are idempotent hermitian and V^* has an orthonormal basis $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_n)$ of eigenvectors of \hat{T}^1 , corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n \in \{0, 1\}$, so

$$[\hat{T}^1]_{\mathbf{f}, \mathbf{f}} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Let us say that this matrix has r entries 1 and the remaining $n - r$ entries 0. By rearranging the eigenvalues and eigenvectors we may assume that the first r entries are 1 and the remaining $n - r$ are 0, i.e.

$$\left([\hat{T}^1]_{\mathbf{f}, \mathbf{f}}\right)_{i,i} = \begin{cases} 1 & : 1 \leq i \leq r \\ 0 & : r + 1 \leq i \leq n. \end{cases}$$

Hence $\hat{T}^1(f_i) = f_i$ for $1 \leq i \leq r$ and $\hat{T}^1(f_i) = 0$ for $r + 1 \leq i \leq n$. Any linear invariant of G is certainly fixed by \hat{T}^1 , so $a_1 \leq r$. On the other hand, by proposition 5.1, $\hat{f}_i := \hat{T}^1(f_i) = \lambda_i f_i$ is an invariant of G for every $1 \leq i \leq r$, so $a_1 \geq r$. Together, $a_1 = r$. \square

Before the final proof, let us introduce a handy notation.

5.7 Definition. Let $p(\lambda) \in \mathbf{C}[\lambda]$ or $p(\lambda) \in \mathbf{C}[[\lambda]]$. Then $[\lambda^i] : p(\lambda)$ denotes the coefficient of λ^i in $p(\lambda)$.

So, for example $[x^2] : 2x^3 + 42x^2 - 6 = 42$ and $[\lambda^d] : \Phi_G(\lambda) = a_d$.

Proof. (Molien's Theorem) We just established the case $d = 1$, so the reader is probably expecting a proof by induction over d . But this is *not* the case. Rather, the case $d = 1$ applies to all $d > 1$. Note that a_d is equal to the number of linearly independent invariants of all of the T_g^d . So Theorem 5.6 gives us

$$\begin{aligned} a_1 &= \text{Tr}(\hat{T}) = \text{Tr}(\hat{T}^1) \quad \text{and} \\ a_d &= \text{Tr}(\hat{T}^d), \end{aligned}$$

where the latter includes the first. From definition 3.1 we also have

$$\hat{T}^1 = \frac{1}{|G|} \sum_{g \in G} T_g^1 \quad \text{and in general} \quad \hat{T}^d = \frac{1}{|G|} \sum_{g \in G} T_g^d,$$

so we already know that

$$a_d = \frac{1}{|G|} \sum_{g \in G} \text{Tr}(T_g^d).$$

So all we need to show is

$$[\lambda^d] : \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{id} - \lambda T_g^1)} = \frac{1}{|G|} \sum_{g \in G} \text{Tr}(T_g^d).$$

We will show that for every summand (group element) the equation

$$[\lambda^d] : \frac{1}{\det(\text{id} - \lambda T_g^1)} = \text{Tr}(T_g^d)$$

holds. From proposition 4.1 we get for every $g \in G$ that

$$\begin{aligned} \text{Tr}(T_g^d) &= \text{Tr}(\text{diag}(\tilde{\omega}_1, \dots, \tilde{\omega}_{\tilde{d}})) \\ &= \tilde{\omega}_1 + \dots + \tilde{\omega}_{\tilde{d}} = \end{aligned}$$

sum of the products of the $\omega_1, \omega_2, \dots, \omega_n$, taken d of them at a time. On the other hand, for the same $g \in G$ we obtain from section 4 that $[T_g^1]_{\mathbf{y}_g^1, \mathbf{y}_g^1} = \text{diag}(\omega_1, \dots, \omega_n)$ so that

$$\begin{aligned} \det(\text{id} - \lambda T_g^1) &= \det(\text{id} - \lambda \cdot \text{diag}(\omega_1, \dots, \omega_n)) \\ &= (1 - \lambda\omega_1)(1 - \lambda\omega_2) \dots (1 - \lambda\omega_n), \end{aligned}$$

so

$$\begin{aligned} \frac{1}{\det(\text{id} - \lambda T_g^1)} &= \frac{1}{(1 - \lambda\omega_1)(1 - \lambda\omega_2) \dots (1 - \lambda\omega_n)} \\ &= \frac{1}{(1 - \lambda\omega_1)} \cdot \frac{1}{(1 - \lambda\omega_2)} \cdot \dots \cdot \frac{1}{(1 - \lambda\omega_n)} \\ &= (1 + \lambda\omega_1 + \lambda^2\omega_1^2 + \dots)(1 + \lambda\omega_2 + \lambda^2\omega_2^2 + \dots) \dots (1 + \lambda\omega_n + \lambda^2\omega_n^2 + \dots) \end{aligned}$$

and here the coefficient of λ^d is also sum of the products of $\omega_1, \omega_2, \dots, \omega_n$, taken d of them at a time.

Again, the last claim

$$\frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{id} - \lambda T_g)} = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{id} - \lambda T_g^1)}$$

follows from the remark at the end of section 3.2, since the sum runs over all group elements. \square

6 Symbol table

a_d	number of linearly independent homogeneous invariants of degree d	R	Big algebra, direct sum of
\tilde{d}	Dimension of R_d	R_d	Direct summand of degree d
\mathcal{B}	ON basis for V	R^G	Ring of invariants of d
G	Finite group	R_d^G	Degree d summand
ω_i	eigenvalue of T_g^1 ([Slo77] = w_i)	T_g	representation of g on V , ([Slo77] $A_\alpha = [T_{g_\alpha}]_{\mathcal{B}, \mathcal{B}}$)
$P(f)$	“Rho” Riesz vector of f .	V	Complex inner product space
ρ	Unitary representation $\rho : G \rightarrow U(V), g \mapsto T_g$	V^*	Algebraic dual of V

7 Lost and found

Some things to explore from here:

- If we know the conjugacy classes of G , we may be able to say more, since every unitary representation splits into irreducible components.
- There seems to be a link to Pólya enumeration.
- We have GAP code, see [GAP].
- An example would be nice.
- Relations on the generators in S of the Cayley graph $\Gamma(G, S)$ should lead to conditions of the minimal polynomial of its adjacency operator $Q(\Gamma(G, S))$.
- Also, Cayley graphs of some finite reflection groups [Hu90] should become accessible.
- Check some more applications, as mentioned in [Slo77].
- For finding invariants, check also [Cox91], Gröbner bases.

References

- [Ant73] Howard Anton, *Elementary Linear Algebra*, 6th ed., John Wiley and Sons, New York, 1973.
- [Bie04] Jürgen Bierbrauer, *Introduction to Coding Theory*, Discrete Mathematics and Its Applications, Volume: 28, CRC Press Inc, Boca Raton, 2004.
- [Cox91] D. Cox, J. Little, D. O’Shea, *Ideals, Varieties, and Algorithms*, Springer-Verlag, New York, 1991.

- [GAP] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.4*; 2004, (<http://www.gap-system.org>).
- [Hu96] John F. Humphreys, *A Course in Group Theory*, Oxford University Press, Oxford, 1994.
- [Hu90] James E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge University Press, Cambridge, 1990.
- [Rom 08] Steven Roman, *Advanced linear algebra, 3rd Edition*, Springer-Verlag, New York, 2008.
- [Sag 91] Bruce E. Sagan, *The Symmetric Group*, Wadsworth & Brooks, Pacific Grove, 1991.
- [Slo77] Neil J. A. Sloane, "Error Correcting Codes and Invariant Theory: New Applications of a Nineteenth-Century Technique", *American Mathematical Monthly*, **84**,(1977), 82–107.
- [Sta79] Richard P. Stanley, "Invariants of Finite Groups and their Applications to Combinatorics", *Bulletin (New Series) of the American Mathematical Society*, **1**, No. 3 (1979), 475–511.
- [Stu93] B. Sturmfels, *Algorithms in Invariant Theory*, Springer-Verlag, Wien, New York, 1993.
- [Tam91] Torbjörn Tambour, *Introduction to Finite Groups and their Representations*, Lecture notes, Lund, 1991.